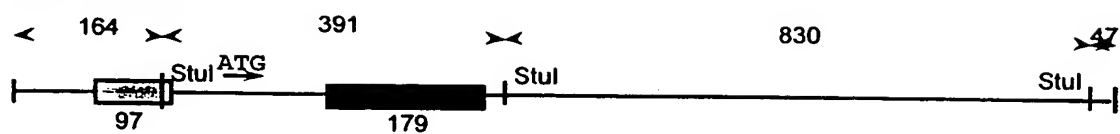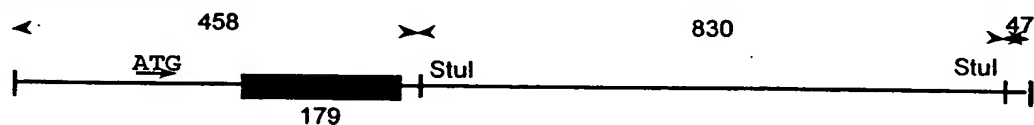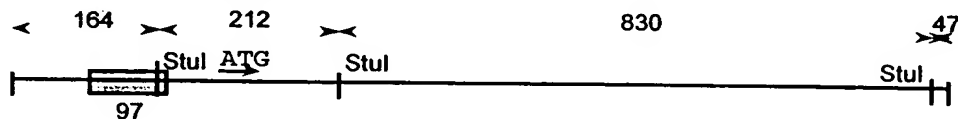Figure 1

**Full Length:**



**Properly Spliced:**



**Downstream Missplice:**



**Unsplice Intron and Downstream Missplice:**



▨ Upstream Splice Site (part of the expression vector)

▨ 'Weird' Splice Site (found in the Gcc cDNA sequence)

Figure 2

<u>r rAG/ gua agu</u>     Consensus Donor


5'- AAG CCG TTG AGT <u>AGG/</u> <u>GTA</u> <u>AG</u>C ATC ATG GCT GGC AGC CTC AC   160-

a) unmodified Gcc


| Lys Pro Leu Ser Arg Val Ser Ile Met Ala Gly Ser Leu Thr   26 |
| --- |


5'- AAG CCG TTG AGT <u>AG</u>A <u>GT</u>C TCC ATC ATG GCT GGC AGC CTC AC   160-
                    *    *  **


b) " modifi4ed Gcc"


        <u>yyy yyy yyy ync ag/G</u>        Consensus Acceptor


5'- TTT CCT GCC <u>CTT</u> GG<u>T</u> A<u>CC</u> <u>TTC</u> <u>AG</u>/C CGC TAT GAG AGT ACA C   340-

a) unmodified Gcc


|    Phe Pro Ala Leu Gly Thr Phe Ser Arg Tyr Glu Ser Thr Arg |
| --- |
| 83 |


5'  TTT CCT GCC <u>C</u>TG GGA A<u>C</u>A <u>TT</u>T TCC CGC TAT GAG AGT ACA C     340-
                *    *   *    *  **


Figure 3 (continued on next page)

b) modified Gcc

Product of Alternatively Spliced mRNA transcript from the
unmodified Gcc cDNA:

NOTE:

Nucleotides associated with RNA splice-sites are underlined.
Those believed to be of most importance to splice site-
recognition are double underlined.

" /" in Consensus Splice Sequence indicates splice site.   " *"
under a particular nucleotide indicates a base changed in the
site-directed mutagenesis procedure.

r = C or A, y = c or u, and n = any.

Figure 3 (continued on next page)

Partial produce of the crypic splicing of Gcc cDNA


5'- AAG CCG TTG AGT AGG CCG CTA TGA GAG TAC AC

| Lys Pro Leu Ser Arg[16] Pro Leu STOP[20]!! |
|---|

**Figure 3 (continued)**

```
              10              20              30              40              50        60
     123456789012345 678901234567890 123456789012345 678901234567890

      NotI
    ▼
    NGCGGCCGCTTAGCT TGACTTAAGAAGGCC GACGCCATGGAGTTT TCAAGTCCTTCCAGA          60
                                        MetGluPhe SerSerProSerArg


    GAGGAATGTCCCAAG CCTTTGAGTAGAGTC TCCATCATGGCTGGC AGCCTCACAGGTTTG          120
    GluGluCysProLys ProLeuSerArgVal SerIleMetAlaGly SerLeuThrGlyLeu


    CTTCTACTTCAGGCA GTGTCGTGGGCATCA GGTGCCCGCCCCTGC ATCCCTAAAAGCTTC          180
    LeuLeuLeuGlnAla ValSerTrpAlaSer GlyAlaArgProCys IleProLysSerPhe


    GGCTACAGCTCGGTG GTGTGTGTCTGCAAT GCCACATACTGTGAC TCCTTTGACCCCCCG          240
    GlyTyrSerSerVal ValCysValCysAsn AlaThrTyrCysAsp SerPheAspProPro


    ACCTTTCCTGCCCTG GGAACATTTTCCCGC TATGAGAGTACACGC AGTGGGCGACGGATG          300
    ThrPheProAlaLeu GlyThrPheSerArg TyrGluSerThrArg SerGlyArgArgMet


    GAGCTGAGTATGGGG CCCATCCAGGCTAAT CACACGGGCACAGGC CTGCTACTGACCCTG          360
    GluLeuSerMetGly ProIleGlnAlaAsn HisThrGlyThrGly LeuLeuLeuThrLeu


    CAGCCAGAACAGAAG TTCCAGAAAGTGAAG GGATTTGGAGGGGCC ATGACAGATGCTGCT          420
    GlnProGluGlnLys PheGlnLysValLys GlyPheGlyGlyAla MetThrAspAlaAla


    GCTCTCAACATCCTT GCCCTGTCACCCCCT GCCCAAAATTTGCTA CTTAAATCGTACTTC          480
    AlaLeuAsnIleLeu AlaLeuSerProPro AlaGlnAsnLeuLeu LeuLysSerTyrPhe


    TCTGAAGAAGGAATC GGATATAACATCATC CGGGTACCCATGGCC AGCTGTGACTTCTCC          540
    SerGluGluGlyIle GlyTyrAsnIleIle ArgValProMetAla SerCysAspPheSer


    ATCCGCACCTACACC TATGCAGACACCCCT GATGATTTCCAGTTG CACAACTTCAGCCTC          600
    IleArgThrTyrThr TyrAlaAspThrPro AspAspPheGlnLeu HisAsnPheSerLeu


    CCAGAGGAAGATACC AAGCTCAAGATACCC CTGATTCACCGAGCC CTGCAGTTGGCCCAG          660
    ProGluGluAspThr LysLeuLysIlePro LeuIleHisArgAla LeuGlnLeuAlaGln


    CGTCCCGTTTCACTC CTTGCCAGCCCCTGG ACATCACCCACTTGG CTCAAGACCAATGGA          720
    ArgProValSerLeu LeuAlaSerProTrp ThrSerProThrTrp LeuLysThrAsnGly
```

Figure 4A(continued on next page)

```
         10              20              30              40              50        60
123456789012345 678901234567890 123456789012345 678901234567890


GCGGTGAATGGGAAG GGGTCACTCAAGGGA CAGCCCGGAGACATC TACCACCAGACCTGG   780
AlaValAsnGlyLys GlySerLeuLysGly GlnProGlyAspIle TyrHisGlnThrTrp


GCCAGATACTTTGTG AAGTTCCTGGATGCC TATGCTGAGCACAAG TTACAGTTCTGGGCA   840
AlaArgTyrPheVal LysPheLeuAspAla TyrAlaGluHisLys LeuGlnPheTrpAla


GTGACAGCTGAAAAT GAGCCTTCTGCTGGG CTGTTGAGTGGATAC CCCTTCCAGTGCCTG   900
ValThrAlaGluAsn GluProSerAlaGly LeuLeuSerGlyTyr ProPheGlnCysLeu


GGCTTCACCCCTGAA CATCAGCGAGACTTA ATTGCCCGTGACCTA GGTCCTACCCTCGCC   960
GlyPheThrProGlu HisGlnArgAspLeu IleAlaArgAspLeu GlyProThrLeuAla


AACAGTACTCACCAC AATGTCCGCCTACTC ATGCTGGATGACCAA CGCTTGCTGCTGCCC   1020
AsnSerThrHisHis AsnValArgLeuLeu MetLeuAspAspGln ArgLeuLeuLeuPro


CACTGGGCAAAGGTG GTACTGACAGACCCA GAAGCAGCTAAATAT GTTCATGGCATTGCT   1080
HisTrpAlaLysVal ValLeuThrAspPro GluAlaAlaLysTyr ValHisGlyIleAla


GTACATTGGTACCTG GACTTTCTGGCTCCA GCCAAAGCCACCCTA GGGGAGACACACCGC   1140
ValHisTrpTyrLeu AspPheLeuAlaPro AlaLysAlaThrLeu GlyGluThrHisArg


CTGTTCCCCAACACC ATGCTCTTTGCCTCA GAGGCCTGTGTGGGC TCCAAGTTCTGGGAG   1200
LeuPheProAsnThr MetLeuPheAlaSer GluAlaCysValGly SerLysPheTrpGlu


CAGAGTGTGCGGCTA GGCTCCTGGGATCGA GGGATGCAGTACAGC CACAGCATCATCACG   1260
GlnSerValArgLeu GlySerTrpAspArg GlyMetGlnTyrSer HisSerIleIleThr


AACCTCCTGTACCAT GTGGTCGGCTGGACC GACTGGAACCTTGCC CTGAACCCCGAAGGA   1320
AsnLeuLeuTyrHis ValValGlyTrpThr AspTrpAsnLeuAla LeuAsnProGluGly


GGACCCAATTGGGTG CGTAACTTTGTCGAC AGTCCCATCATTGTA GACATCACCAAGGAC   1380
GlyProAsnTrpVal ArgAsnPheValAsp SerProIleIleVal AspIleThrLysAsp


ACGTTTTACAAACAG CCCATGTTCTACCAC CTTGGCCATTTCAGC AAGTTCATTCCTGAG   1440
ThrPheTyrLysGln ProMetPheTyrHis LeuGlyHisPheSer LysPheIleProGlu
```
Figure 4A(continued on next page)

```
                10              20              30              40              50              60
        1234567890123456 789012345678901234 567890123456789 012345678901234567890
```

GGCTCCCAGAGAGTG GGGCTGGTTGCCAGT CAGAAGAACGACCTG GACGCAGTGGCATTG      1500
GlySerGlnArgVal GlyLeuValAlaSer GlnLysAsnAspLeu AspAlaValAlaLeu


ATGCATCCCGATGGC TCTGCTGTTGTGGTC GTGCTAAACCGCTCC TCTAAGGATGTGCCT      1560
MetHisProAspGly SerAlaValValVal ValLeuAsnArgSer SerLysAspValPro


CTTACCATCAAGGAT CCTGCTGTGGGCTTC CTGGAGACAATCTCA CCTGGCTACTCCATT      1620
LeuThrIleLysAsp ProAlaValGlyPhe LeuGluThrIleSer ProGlyTyrSerIle


CACACCTACCTGTGG CATCGCCAGTGATGG AGCAGATACTCAAGG AGGCACTGGGCTCAG      1680
HisThrTyrLeuTrp HisArgGln


                                                                    NotI
CCTGGGCATTAAAGG GACAGAGTCAGCGAA TTCTGCAGATATCCA TCACACTGGCGGCCG      1740


C                                                                                1741


Figure 4A(continued)

```
        10            20            30            40            50            60
123456789012345 678901234567890 123456789012345 678901234567890

NotI
▼
NGCGGCCGCTTAGCT TGACTTAAGAAGGCC GACGCCATGGAGTTT TCAAGTCCTTCCAGA        60
                                      MetGluPhe SerSerProSerArg


GAGGAATGTCCCAAG CCTTTGAGTAGAGTC TCCATCATGGCTGGC AGCCTCACAGGTTTG       120
GluGluCysProLys ProLeuSerArgVal SerIleMetAlaGly SerLeuThrGlyLeu


CTTCTACTTCAGGCA GTGTCGTGGGCATCA GGTGCCCGCCCCTGC ATCCCTAAAAGCTTC       180
LeuLeuLeuGlnAla ValSerTrpAlaSer GlyAlaArgProCys IleProLysSerPhe


GGCTACAGCTCGGTG GTGTGTGTCTGCAAT GCCACATACTGTGAC TCCTTTGACCCCCCG       240
GlyTyrSerSerVal ValCysValCysAsn AlaThrTyrCysAsp SerPheAspProPro


ACCTTTCCTGCCCTG GGAACATTTTCCCGC TATGAGAGTACACGC AGTGGGCGACGGATG       300
ThrPheProAlaLeu GlyThrPheSerArg TyrGluSerThrArg SerGlyArgArgMet


GAGCTGAGTATGGGG CCCATCCAGGCTAAT CACACGGGCACAGGC CTGCTACTGACCCTG       360
GluLeuSerMetGly ProIleGlnAlaAsn HisThrGlyThrGly LeuLeuLeuThrLeu


CAGCCAGAACAGAAG TTCCAGAAAGTGAAG GGATTTGGAGGGGCC ATGACAGATGCTGCT       420
GlnProGluGlnLys PheGlnLysValLys GlyPheGlyGlyAla MetThrAspAlaAla


GCTCTCAACATCCTT GCCCTGTCACCCCCT GCCCAAAATTTGCTA CTTAAATCGTACTTC       480
AlaLeuAsnIleLeu AlaLeuSerProPro AlaGlnAsnLeuLeu LeuLysSerTyrPhe


TCTGAAGAAGGAATC GGATATAACATCATC CGGGTACCCATGGCC AGCTGTGACTTCTCC       540
SerGluGluGlyIle GlyTyrAsnIleIle ArgValProMetAla SerCysAspPheSer


ATCCGCACCTACACC TATGCAGACACCCCT GATGATTTCCAGTTG CACAACTTCAGCCTC       600
IleArgThrTyrThr TyrAlaAspThrPro AspAspPheGlnLeu HisAsnPheSerLeu


CCAGAGGAAGATACC AAGCTCAAGATACCC CTGATTCACCGAGCC CTGCAGTTGGCCCAG       660
ProGluGluAspThr LysLeuLysIlePro LeuIleHisArgAla LeuGlnLeuAlaGln


CGTCCCGTTTCACTC CTTGCCAGCCCCTGG ACATCACCCACTTGG CTCAAGACCAATGGA       720
ArgProValSerLeu LeuAlaSerProTrp ThrSerProThrTrp LeuLysThrAsnGly
```

**Mutations made to the Gcc insert are underlined**

Figure 4B (continued on next page)

```
            10              20              30              40              50            60
1234567890123 45 678901234567890 123456789012345 678901234567890

GCGGTGAATGGGAAG GGGTCACTCAAGGGA CAGCCCGGAGACATC TACCACCAGACCTGG   780
AlaValAsnGlyLys GlySerLeuLysGly GlnProGlyAspIle TyrHisGlnThrTrp

GCCAGATACTTTGTG AAGTTCCTGGATGCC TATGCTGAGCACAAG TTACAGTTCTGGGCA   840
AlaArgTyrPheVal LysPheLeuAspAla TyrAlaGluHisLys LeuGlnPheTrpAla

GTGACAGCTGAAAAT GAGCCTTCTGCTGGG CTGTTGAGTGGATAC CCCTTCCAGTGCCTG   900
ValThrAlaGluAsn GluProSerAlaGly LeuLeuSerGlyTyr ProPheGlnCysLeu

GGCTTCACCCCTGAA CATCAGCGAGACTTA ATTGCCCGTGACCTA GGTCCTACCCTCGCC   960
GlyPheThrProGlu HisGlnArgAspLeu IleAlaArgAspLeu GlyProThrLeuAla

AACAGTACTCACCAC AATGTCCGCCTACTC ATGCTGGATGACCAA CGCTTGCTGCTGCCC   1020
AsnSerThrHisHis AsnValArgLeuLeu MetLeuAspAspGln ArgLeuLeuLeuPro

CACTGGGCAAAGGTG GTACTGACAGACCCA GAAGCAGCTAAATAT GTTCATGGCATTGCT   1080
HisTrpAlaLysVal ValLeuThrAspPro GluAlaAlaLysTyr ValHisGlyIleAla

GTACATTGGTACCTG GACTTTCTGGCTCCA GCCAAAGCCACCCTA GGGGAGACACACCGC   1140
ValHisTrpTyrLeu AspPheLeuAlaPro AlaLysAlaThrLeu GlyGluThrHisArg

CTGTTCCCCAACACC ATGCTCTTTGCCTCA GAGGCCTGTGTGGGC TCCAAGTTCTGGGAG   1200
LeuPheProAsnThr MetLeuPheAlaSer GluAlaCysValGly SerLysPheTrpGlu

CAGAGTGTGCGGCTA GGCTCCTGGGATCGA GGGATGCAGTACAGC CACAGCATCATCACG   1260
GlnSerValArgLeu GlySerTrpAspArg GlyMetGlnTyrSer HisSerIleIleThr

AACCTCCTGTACCAT GTGGTCGGCTGGACC GACTGGAACCTTGCC CTGAACCCCGAAGGA   1320
AsnLeuLeuTyrHis ValValGlyTrpThr AspTrpAsnLeuAla LeuAsnProGluGly

GGACCCAATTGGGTG CGTAACTTTGTCGAC AGTCCCATCATTGTA GACATCACCAAGGAC   1380
GlyProAsnTrpVal ArgAsnPheValAsp SerProIleIleVal AspIleThrLysAsp

ACGTTTTACAAACAG CCCATGTTCTACCAC CTTGGCCATTTCAGC AAGTTCATTCCTGAG   1440
ThrPheTyrLysGln ProMetPheTyrHis LeuGlyHisPheSer LysPheIleProGlu
```

Mutations made to the Gcc insert are underlined

Figure 4B (continued on next page)

```
         10            20             30            40             50        60
123456789012345 678901234567890 123456789012345 678901234567890


GGCTCCCAGAGAGTG GGGCTGGTTGCCAGT CAGAAGAACGACCTG GACGCAGTGGCATTG       1500
GlySerGlnArgVal GlyLeuValAlaSer GlnLysAsnAspLeu AspAlaValAlaLeu


ATGCATCCCGATGGC TCTGCTGTTGTGGTC GTGCTAAACCGCTCC TCTAAGGATGTGCCT       1560
MetHisProAspGly SerAlaValValVal ValLeuAsnArgSer SerLysAspValPro


CTTACCATCAAGGAT CCTGCTGTGGGCTTC CTGGAGACAATCTCA CCTGGCTACTCCATT       1620
LeuThrIleLysAsp ProAlaValGlyPhe LeuGluThrIleSer ProGlyTyrSerIle


CACACCTACCTGTGG CATCGCCAGTGATGG AGCAGATACTCAAGG AGGCACTGGGCTCAG       1680
HisThrTyrLeuTrp HisArgGln


                                                        NotI
CCTGGGCATTAAAGG GACAGAGTCAGCGAA TTCTGCAGATATCCA TCACACTGGCGGCCG       1740


C                                                                      1741
```

Mutations made to the Gcc insert are underlined

Figure 4B (continued)

```
              10          20          30          40          50          60
    123456789012345 678901234567890 123456789012345 678901234567890


    GGCTCCCAGAGAGTG GGGCTGGTTGCCAGT CAGAAGAACGACCTG GACGCAGTGGCATTG    1500
    GlySerGlnArgVal GlyLeuValAlaSer GlnLysAsnAspLeu AspAlaValAlaLeu


    ATGCATCCCGATGGC TCTGCTGTTGTGGTC GTGCTAAACCGCTCC TCTAAGGATGTGCCT    1560
    MetHisProAspGly SerAlaValValVal ValLeuAsnArgSer SerLysAspValPro


    CTTACCATCAAGGAT CCTGCTGTGGGCTTC CTGGAGACAATCTCA CCTGGCTACTCCATT    1620
    LeuThrIleLysAsp ProAlaValGlyPhe LeuGluThrIleSer ProGlyTyrSerIle


    CACACCTACCTGTGG CATCGCCAGTGATGG AGCAGATACTCAAGG AGGCACTGGGCTCAG    1680
    HisThrTyrLeuTrp HisArgGln

                                                              NotI
    CCTGGGCATTAAAGG GACAGAGTCAGCGAA TTCTGCAGATATCCA TCACACTGGCGGCCG    1740


    C                                                                 1741
```

Mutations made to the Gcc insert are underlined

Figure 4B (continued)